MULTISEQUENCE DATA REPRESENTATION
ABSTRACT

Genetic sequence data occurring in genome sequences is represented for efficient access

5   of the sequence information in a defined storage scheme. A described replet-sequence matrix data structure allows the compression and efficient access of sequence information. The data structure allows the dynamic change of ontology: the replet-information table can evolve by adding, updating, removing replets, and the set of replets present in the table represent the ontology at the moment. The data structure enables the

10   sequence information to be processed in parallel, and also enables multiple views of the sequence data to exist along with replet specific information.

15